

# Python: module cdms.cdurlparse

## *cdms.cdurlparse*

[index](#)

Parse (absolute and relative) URLs. See [RFC 1808](#): "Relative Uniform Resource Locators", by R. Fielding, UC Irvine, June 1995.

### *Modules*

[string](#)

### *Functions*

#### ***clear\_cache()***

Clear the parse cache.

#### ***test()***

#### ***urldefrag(url)***

Removes any existing fragment from URL.

Returns a tuple of the defragmented URL and the fragment. If the URL contained no fragments, the second element is the empty string.

#### ***urljoin(base, url, allow\_fragments=1)***

# Join a base URL and a possibly relative URL to form an absolute  
# interpretation of the latter.

#### ***urlparse(url, scheme="", allow\_fragments=1)***

# Parse a URL into 6 components:  
# <scheme>://<netloc>/<path>;<params>?<query>#<fragment>  
# Return a 6-tuple: (scheme, netloc, path, params, query, fragment)  
# Note that we don't break the components up in smaller bits  
# (e.g. netloc is a single string) and we don't expand % escapes.

#### ***urlunparse((scheme, netloc, url, params, query, fragment))***

# Put a parsed URL back together again. This may result in a slight  
# different, but equivalent URL, if the URL that was parsed originally  
# had redundant delimiters, e.g. a ? with an empty query (the draft  
# states that these are equivalent).

### *Data*

```
MAX_CACHE_SIZE = 20
non_hierarchical = ['gopher', 'hdl', 'mailto', 'news', 'telnet', 'wais', 'snews']
scheme_chars =
'abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789+-.'
test_input = '\n http://a/b/c/d\n\n g:h = <URL:g... http:g?y./x = <URL:http://a/b/c/g?y./x>\n'
uses_fragment = ['ftp', 'hdl', 'http', 'ldap', 'gopher', 'news', 'nntp', 'wais', 'https', 'shttp', 'snews',
'file', 'prospero', ""]
uses_netloc = ['ftp', 'http', 'ldap', 'gopher', 'nntp', 'telnet', 'wais', 'file', 'https', 'shttp', 'snews',
'prospero', ""]
uses_params = ['ftp', 'hdl', 'prospero', 'http', 'ldap', 'https', 'shttp', ""]
uses_query = ['http', 'ldap', 'wais', 'https', 'shttp', 'gopher', ""]
uses_relative = ['ftp', 'http', 'ldap', 'gopher', 'nntp', 'wais', 'file', 'https', 'shttp', 'prospero', ""]
```